

Abstract

Method for forming and/or updating dictionaries for the automatic reading of addresses

The reading results of an agreed number of images of items, achieved by the OCR reader, are temporarily stored subdivided into reading results which are read unambiguously and reading results which are rejected.

Then classes of words or word groups belonging together of the reading results temporarily stored and rejected, consisting in each case of n address words, $n = 1, 2, \dots, a$, with interword gaps m , $m = 0, 1, \dots, b$ are formed which do not drop below a particular similarity factor referred to in each case a particular n and m value between them. In the dictionary or dictionaries of the associated address areas, representatives, at least, of the classes whose frequency exceeds a predetermined value are included.

Figure 1